Combinatorics, Probability and Computing (2012) 21, 952–962.
doi:10.1017/S0963548312000235

Dense Graphs With a Large Triangle Cover Have a Large Triangle Packing

RAPHAEL YUSTER

Department of Mathematics, University of Haifa, Haifa 31905, Israel
(e-mail: raphy@math.haifa.ac.il)

Received 3 September 2010; revised 20 April 2012

It is well known that a graph with m edges can be made triangle-free by removing (slightly less than) $m/2$ edges. On the other hand, there are many classes of graphs which are hard to make triangle-free, in the sense that it is *necessary* to remove roughly $m/2$ edges in order to eliminate all triangles.

We prove that dense graphs that are hard to make triangle-free have a large packing of pairwise edge-disjoint triangles. In particular, they have more than $m(1/4 + c\beta)$ pairwise edge-disjoint triangles where β is the density of the graph and $c \geqslant \frac{1}{100}$ is an absolute constant. This improves upon a previous $m(1/4 - o(1))$ bound which follows from the asymptotic validity of Tuza's conjecture for dense graphs. We conjecture that such graphs have an asymptotically optimal triangle packing of size $m(1/3 - o(1))$.

We extend our result from triangles to larger cliques and odd cycles.

AMS 2010 *Mathematics subject classification*: Primary 05C70

1. Introduction

All graphs in this paper are finite, undirected, and simple. A *triangle edge cover* in a graph is a set of edges meeting all triangles. In other words, the removal of a triangle edge cover results in a triangle-free graph. Dually, a *triangle packing* in a graph is a set of pairwise edge-disjoint triangles. We denote by $\tau_3(G)$ the minimum size of a triangle edge cover and by $\nu_3(G)$ the maximum size of a triangle packing of a graph G. It is easily observed that

$$\nu_3(G) \leqslant \tau_3(G) \leqslant 3\nu_3(G).$$

The first inequality follows from the fact that one must delete at least one edge from each triangle in a triangle packing in order to obtain a triangle-free graph. The second inequality follows from the fact that deleting all edges of all triangles in a maximum triangle packing results in a triangle-free graph. A long-standing conjecture of Tuza [6] states that this second inequality is not optimal.

Conjecture 1.1 (Tuza [6]). $\tau_3(G) \leqslant 2\nu_3(G)$.

This conjecture, if true, is best possible, as can be seen by taking, say, $G = K_4$ or $G = K_5$. The best upper bound for $\tau_3(G)$ is due to Haxell [2], who proved that $\tau_3(G) \leqslant (3 - \frac{3}{23})\nu_3(G)$.

However, there is an important setting where, asymptotically, Tuza's conjecture holds. This is the *dense graph* setting. To state this result we need first to consider the fractional relaxations of $\tau_3(G)$ and $\nu_3(G)$. A *fractional triangle edge cover* assigns non-negative weights to the edges so that the resulting weight of each triangle (being the sum of the weights of its edges) is at least 1. Dually, a *fractional triangle packing* assigns non-negative weights to the triangles so that the resulting weight of each edge (being the sum of the weights of the triangles it meets) is at most 1. The goal is thus to minimize the sum of the weights of a fractional triangle edge cover and to maximize the sum of the weights of a fractional triangle packing. Therefore let $\tau_3^*(G)$ and $\nu_3^*(G)$ be the fractional relaxations of $\tau_3(G)$ and $\nu_3(G)$ respectively. By linear programming duality we have $\tau_3^*(G) = \nu_3^*(G)$, and, trivially, $\tau_3(G) \geqslant \tau_3^*(G)$ and $\nu_3(G) \leqslant \nu_3^*(G)$. Krivelevich [5] proved that Tuza's conjecture holds in a mixed fractional-integral setting. Namely, he proved the following result.

Theorem 1.2 (Krivelevich [5]). *For any graph G we have $\tau_3(G) \leqslant 2\nu_3^*(G)$ and $\tau_3^*(G) \leqslant 2\nu_3(G)$.*

The inequality $\tau_3^*(G) \leqslant 2\nu_3(G)$ is tight (*e.g.*, K_4) and the inequality $\tau_3(G) \leqslant 2\nu_3^*(G)$ is known to be asymptotically tight. A few years later, Haxell and Rödl [3] (see also [7]) proved that $|\nu_3(G) - \nu_3^*(G)| = o(n^2)$ for n-vertex graphs G. In other words, in graphs that contain a quadratic number of pairwise edge-disjoint triangles, $\nu_3(G)$ and $\nu_3^*(G)$ are asymptotically the same. These results imply the following theorem.

Theorem 1.3. $\tau_3(G) \leqslant 2\nu_3(G) + o(n^2)$.

In light of the fact that Tuza's conjecture is optimal, it is interesting to ask whether the constant 2 in Theorem 1.3 is also optimal (notice that this question becomes non-trivial for dense graphs with $\tau_3(G) = \Theta(n^2)$). Perhaps the most interesting case to consider is when $\tau_3(G)$ is as large as one can expect it to be.

It is well known that every graph with m edges can be made bipartite by removing from it less than $m/2$ edges (see [1] for the tightest known bounds). In particular, $\tau_3(G) \leqslant m/2 - o(m)$. On the other hand, there are many different types of graphs that are hard to make triangle-free, that is, graphs for which $\tau_3(G) \geqslant m/2 - o(m)$. For example, complete graphs are hard to make triangle-free, and (sufficiently dense) random graphs are hard to make triangle-free. It is also easy to construct many other families of graphs that are hard to make triangle-free. Let us formalize this notion. We say that a graph G is $(1-\delta)$-*hard to make* Δ-*free* if $\tau_3(G) \geqslant (1-\delta)(m/2)$.

The following is an immediate consequence of Theorem 1.3.

Corollary 1.4. *Let G be a graph with m edges that is $(1 - o_n(1))$-hard to make Δ-free. Then,*

$$\nu_3(G) \geqslant \frac{m}{4} - o(n^2).$$

We conjecture that Corollary 1.4 is *not* optimal, and that $m/4$ can be replaced with $m/3$. Formally, we conjecture the following.

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

Conjecture 1.5. *For every $\epsilon > 0$ and $\beta > 0$ there exist $N = N(\epsilon, \beta)$ and $\delta = \delta(\epsilon, \beta)$ such that, for all graphs with $n > N$ vertices and with $m \geqslant \beta n^2$ edges that are $(1-\delta)$-hard to make Δ-free,*

$$\nu_3(G) \geqslant (1-\epsilon)\frac{m}{3}.$$

Since for any m-edge graph we have $\nu_3(G) \leqslant m/3$, Conjecture 1.5 states that dense graphs that are hard to make Δ-free have an asymptotically optimal triangle packing: all but a negligible fraction of the edges are packed.

A weakened, but still challenging version of Conjecture 1.5 asks for a constant improvement over the $1/4$ bound in Corollary 1.4.

Conjecture 1.6. *There exists $\alpha > 0$ so that for all $\beta > 0$ there exist $N = N(\beta)$ and $\delta = \delta(\beta)$ such that, for all graphs with $n > N$ vertices and with $m \geqslant \beta n^2$ edges that are $(1-\delta)$-hard to make Δ-free,*

$$\nu_3(G) \geqslant (1+\alpha)\frac{m}{4}.$$

A further weakening of Conjecture 1.6 allows the improvement α to depend on the density β. The main result of this paper proves that such an improvement always exists. Hence, for any fixed density, our main result shows that the constant $1/4$ in Corollary 1.4 is not optimal, and can be replaced with a larger constant.

Theorem 1.7. *For every $\beta > 0$ there are $N = N(\beta)$ and $\delta = \delta(\beta)$ such that, for all graphs with $n > N$ vertices and with $m \geqslant \beta n^2$ edges that are $(1-\delta)$-hard to make Δ-free,*

$$\nu_3(G) \geqslant \left(1 + \frac{\beta}{100}\right)\frac{m}{4}.$$

The constant 100 in Theorem 1.7 is by no means optimal and it can be somewhat reduced at the price of complicating the calculations. Since this has no qualitative impact on the statement of Theorem 1.7, we make no effort to optimize it. We also note that $\delta = \delta(\beta)$ is a moderate function. As shown in the proof, it suffices to take $\delta = \beta^2/99$.

The next section contains the proof of Theorem 1.7. Section 3 considers larger cliques. We prove a bound for the edge-covering number of K_k in terms of the fractional edge-covering number of K_k, and then use it to extend Theorem 1.7 to larger cliques. Section 4 contains some concluding remarks: a sketch of a generalization of Theorem 1.7 to larger odd cycles, and an improved integrality gap for the problem of determining the 'maximal triangle-free subgraph' in dense graphs.

2. Packing triangles in graphs that are hard to make triangle-free

Since $\nu_3^*(G) = \tau_3^*(G)$, and since by the result of Haxell and Rödl mentioned earlier we have $\nu_3^*(G) \leqslant \nu_3(G) + o(n^2)$, the following theorem immediately implies Theorem 1.7.

Theorem 2.1. *For every $\beta > 0$ there exists an integer $N = N(\beta)$ such that, for all graphs with $n > N$ vertices and with $m \geqslant \beta n^2$ edges that are $(1-\beta^2/99)$-hard to make Δ-free,*

$$\tau_3^*(G) \geqslant \left(1 + \frac{\beta}{99}\right)\frac{m}{4}.$$

We will therefore prove Theorem 2.1, and hence obtain a proof for Theorem 1.7 as well.

We first need to recall some known facts from linear programming. For a graph G, let $E(G)$ and $T(G)$ denote the sets of edges and triangles of G, respectively. Let $f : E(G) \to [0,1]$ be a minimum fractional triangle edge cover so that $\sum_{e\in E(G)} f(e) = \tau_3^*(G)$, and let $g : T(G) \to [0,1]$ be a maximum fractional triangle packing so that $\sum_{t\in T(G)} g(t) = \nu_3^*(G)$. Then, the duality theorem of linear programming states that $\tau_3^*(G) = \nu_3^*(G)$ and (one of) the complementary slackness conditions states that

$$f(e) > 0 \text{ implies } \sum_{t\ni e} g(t) = 1. \tag{2.1}$$

We designate two sets of edges.

- Let $F_0 \subset E(G)$ be $F_0 = \{e \mid f(e) = 0\}$.
- Let $F_1 \subset E(G)$ be $F_1 = \{e \mid f(e) = 1\}$.

The proof of Theorem 2.1 is split into three cases, according to the cardinalities of F_0 and F_1. The first two cases are easy. The first is when F_1 is relatively large and the second is when F_0 is relatively small. The remaining case, where F_1 is relatively small and F_0 is relatively large, is more difficult. It will be convenient to assume, without loss of generality, that $m = \beta n^2$. Observe that this immediately implies the proof for $m \geqslant \beta n^2$. We also set $\delta = \beta^2/99$ so that the assumption in Theorem 2.1 is that the graph is $(1-\delta)$-hard to make Δ-free.

Case 1: $|F_1| > (\delta + \frac{\beta}{99})m/2$. Define $G_1 = G(V, E \setminus F_1)$ to be the graph obtained from G by deleting the edges having weight 1. We observe that

$$\tau_3(G_1) \geqslant \tau_3(G) - |F_1|, \tag{2.2}$$
$$\tau_3^*(G_1) \leqslant \tau_3^*(G) - |F_1|, \tag{2.3}$$
$$\tau_3^*(G_1) \geqslant \frac{1}{2}\tau_3(G_1). \tag{2.4}$$

Indeed, (2.2) holds since we have deleted $|F_1|$ edges, (2.3) holds since the total deleted weight is $|F_1|$, and (2.4) holds by Theorem 1.2. Using these inequalities and the assumption on the size of F_1, we have

$$\begin{aligned}\tau_3^*(G) &\geqslant \tau_3^*(G_1) + |F_1| \\ &\geqslant \frac{1}{2}\tau_3(G_1) + |F_1| \\ &\geqslant \frac{1}{2}(\tau_3(G) - |F_1|) + |F_1| \\ &= \frac{1}{2}\tau_3(G) + \frac{1}{2}|F_1|\end{aligned}$$

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

$$\geqslant \frac{1}{2}(1-\delta)\frac{m}{2} + \left(\delta + \frac{\beta}{99}\right)\frac{m}{4}$$
$$= \left(1 + \frac{\beta}{99}\right)\frac{m}{4}.$$

Case 2: $|F_0| < (1 - \frac{3\beta}{99})m/4$. The complementary slackness condition (2.1) implies that

$$\sum_{e \in E \setminus F_0} \sum_{t \ni e} g(t) = m - |F_0|.$$

As each triangle is counted at most three times, we have that

$$\tau_3^*(G) = v_3^*(G) \geqslant \frac{1}{3}(m - |F_0|).$$

Using the assumption on the size of F_0, we obtain

$$\tau_3^*(G) \geqslant \frac{1}{3}\left(m - \left(1 - \frac{3\beta}{99}\right)m/4\right) = \left(1 + \frac{\beta}{99}\right)\frac{m}{4}.$$

Case 3: $|F_0| \geqslant (1 - \frac{3\beta}{99})m/4$ and $|F_1| \leqslant (\delta + \frac{\beta}{99})m/2$. Since $\beta \leqslant 1/2$, our assumptions in this case imply in particular that

$$|F_0| \geqslant 0.246m, \tag{2.5}$$
$$|F_1| \leqslant \beta m/99. \tag{2.6}$$

For a subset of vertices $A \subset V$, let $E(A)$ denote the set of edges of G with both endpoints in A. If A and B are disjoint subsets of vertices, then $E(A,B)$ denotes the set of edges with one endpoint in A and the other in B.

Lemma 2.2. *G contains two disjoint subsets of vertices A and B such that*

$$|E(A,B)| - |E(A)| - |E(B)| > \frac{\beta^2 m}{99}.$$

Proof. Consider the graph $H = G(V, F_0)$ consisting only of the edges of G having weight zero. Notice that H is still dense as it has at least $0.246m$ edges, and that H is triangle-free since otherwise f would not have been a fractional triangle edge cover.

Consider a random subset C of $c = \lceil 1/(0.246\beta) \rceil$ vertices. We say that a vertex x is *dominated* by C if some vertex of C is a neighbour of x in H. Clearly, the probability that a vertex is *not* dominated by C is less than $(1 - c/n)^{d_x}$, where d_x is the degree of x in H. Let Q denote the set of all edges of H that are incident with vertices that are not dominated by C. Hence, the expected size of Q satisfies

$$\mathbb{E}[|Q|] < \sum_{x \in V} d_x \left(1 - \frac{c}{n}\right)^{d_x} \leqslant \sum_{x \in V} d_x e^{-d_x c/n} \leqslant \sum_{x \in V} \frac{n}{ec} = \frac{n^2}{ec}.$$

In particular, there exists a choice of C such that after removing from H the vertices that are not dominated by C, we remain with a subgraph H' whose number of edges is at least

$$0.246m - \frac{n^2}{ec} \geqslant 0.246m - \frac{0.246\beta n^2}{e} \geqslant 0.246m - \frac{0.246m}{e} = 0.246(1-1/e)m.$$

As each vertex of H' is dominated by C, let us select for each vertex of H' a vertex of C that dominates it. This partitions the vertices of H' into c parts $\{A_u \mid u \in C\}$, where A_u consists of the vertices of H' that chose u as their dominating vertex. Observe also that each A_u induces an independent set in H, since otherwise we would have, together with u, a triangle in H, contradicting its triangle-freeness. As $\cup_{u \neq v \in C} E(A_u, A_v)$ contains all the edges of H' and no edge is counted more than once, we have

$$\sum_{u \neq v \in C} |E(A_u, A_v)| \geqslant 0.246(1-1/e)m. \tag{2.7}$$

Moreover, $E(A_u)$ contains only edges of F_1. Indeed, if we had an edge $(a, a') \in E(A_u)$ with $f((a, a')) < 1$, then the triangle (a, a', u) would have weight less than 1 (the edges (u, a) and (u, a') have weight 0 as they both belong to F_0). But a triangle cannot have weight less than 1 since f is a fractional triangle edge cover. It follows that

$$\sum_{u \in C} |E(A_u)| \leqslant |F_1|. \tag{2.8}$$

Using (2.7) and (2.8) we have

$$\sum_{u \neq v \in C} [|E(A_u, A_v)| - |E(A_u)| - |E(A_v)|] \geqslant 0.246(1-1/e)m - (c-1)|F_1|.$$

It follows that there is a particular choice of pair A_u, A_v for which

$$|E(A_u, A_v)| - |E(A_u)| - |E(A_v)| \geqslant \frac{1}{\binom{c}{2}} 0.246(1-1/e)m - \frac{2}{c}|F_1| \geqslant \frac{\beta^2 m}{54} - \frac{\beta^2 m}{200} > \frac{\beta^2 m}{99},$$

where we have used (2.6) and $c = \lceil 1/(0.246\beta) \rceil$. □

Let A and B be disjoint subsets satisfying Lemma 2.2. To complete the proof of Case 3, we proceed as follows. We split the vertices of $V \setminus (A \cup B)$ into two parts X and Y at random. We consider the cut $(A \cup X, B \cup Y)$ and compute the expected number of edges crossing it. Each edge of $E(A, B)$ crosses it by definition. On the other hand, each edge with at least one endpoint in $X \cup Y$ crosses it with probability $1/2$. The expected size of this cut is therefore

$$|E(A,B)| + \frac{1}{2}(m - |E(A,B)| - |E(A)| - |E(B)|) > \frac{m}{2} + \frac{\beta^2 m}{198}.$$

Hence, such a cut exists, implying that we can remove from G the non-edges of this cut to obtain a triangle-free (in fact, bipartite) subgraph. The number of edges thus removed is less than

$$m - \frac{m}{2} - \frac{m\beta^2}{198} = \frac{m}{2}\left(1 - \frac{\beta^2}{99}\right) = \frac{m}{2}(1-\delta),$$

contradicting the assumption that $\tau_3(G) \geqslant \frac{m}{2}(1-\delta)$. This completes the proof of Theorem 2.1, which, as noted earlier, implies Theorem 1.7. □

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

3. Larger cliques

Throughout this section we fix $k \geqslant 4$, we let $\tau_k(G)$ denote the minimum size of a K_k edge cover, and let $\nu_k(G)$ denote the maximum size of a K_k-packing of a graph G. The trivial bounds in this case are $\nu_k(G) \leqslant \tau_k(G) \leqslant \binom{k}{2}\nu_k(G)$.

Denoting by $\tau_k^*(G)$ and $\nu_k^*(G)$ the respective (and equal) fractional parameters, Krivelevich's proof for triangles [5] can be generalized to yield

$$\tau_k(G) \leqslant \left(\binom{k}{2} - 1\right)\tau_k^*(G). \tag{3.1}$$

We omit the details of this easy generalization since the bounds we shall obtain in this section are better.

As for the case of triangles, the theorem of Haxell and Rödl [3] asserts that $|\nu_k(G) - \nu_k^*(G)| = o(n^2)$. Thus, an immediate corollary analogous to Theorem 1.3 is as follows.

Corollary 3.1. $\tau_k(G) \leqslant (\binom{k}{2} - 1)\nu_k(G) + o(n^2)$.

The goal of this section is to prove a significantly better bound, replacing $\binom{k}{2} - 1$ with a much smaller value. We shall do that by improving upon (3.1).

Theorem 3.2. $\tau_k(G) \leqslant \lfloor k^2/4 \rfloor \tau_k^*(G)$.

Proof. Consider the following process which creates a sequence of spanning subgraphs G_i of G, starting with $G = G_0$. Each G_i is obtained from its predecessor G_{i-1} by deleting a single edge according to the rule specified below. We will halt this process once the rule cannot be applied. We denote the final graph in our sequence by G_t. Hence we have $0 \leqslant t \leqslant m$.

Let f_i and g_i be a minimum fractional K_k edge cover and a maximum fractional K_k-packing of G_i, respectively. Assume first that some K_k of G_i contains $\binom{k}{2} - \lfloor k^2/4 \rfloor$ edges that are assigned weight 0 by f_i. This means that the total weight of the remaining $\lfloor k^2/4 \rfloor$ edges of this K_k is at least 1, so there is some edge e_i with $f_i(e_i) \geqslant 1/\lfloor k^2/4 \rfloor$. We let $G_{i+1} = G_i - \{e_i\}$. If no K_k of G_i contains $\binom{k}{2} - \lfloor k^2/4 \rfloor$ edges that are assigned weight 0 by f_i, then we halt the sequence and $G_i = G_t$ is the final graph in the sequence.

We observe the following inequalities:

$$\tau_k(G_t) \geqslant \tau_k(G) - t, \tag{3.2}$$

$$\tau_k^*(G_t) \leqslant \tau_k^*(G) - \frac{t}{\lfloor k^2/4 \rfloor}, \tag{3.3}$$

$$\tau_k^*(G_t) \geqslant \frac{\tau_k(G_t)}{\binom{k}{2}}. \tag{3.4}$$

Indeed, (3.2) holds since we have deleted t edges to get from G to G_t, (3.3) holds since

$$\tau_k^*(G_{i+1}) \leqslant \tau_k^*(G_i) - 1/\lfloor k^2/4 \rfloor,$$

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

and (3.4) is the trivial bound. Using these inequalities we have

$$\begin{aligned}\tau_k^*(G) &\geqslant \tau_k^*(G_t) + \frac{t}{\lfloor k^2/4 \rfloor} \\ &\geqslant \frac{\tau_k(G_t)}{\binom{k}{2}} + \frac{t}{\lfloor k^2/4 \rfloor} \\ &\geqslant \frac{\tau_k(G) - t}{\binom{k}{2}} + \frac{t}{\lfloor k^2/4 \rfloor} \\ &= \frac{\tau_k(G)}{\binom{k}{2}} - \frac{t}{\binom{k}{2}} + \frac{t}{\lfloor k^2/4 \rfloor}. \end{aligned} \tag{3.5}$$

Let $0 \leqslant \alpha \leqslant 1$ be chosen such that G_t has $\alpha(m-t)$ edges that are assigned weight 0 by f_t. Thus, $(1-\alpha)(m-t)$ edges of G_t are assigned positive weight and, using complementary slackness as in (2.1), we obtain

$$\sum_{e: f_i(e)>0} \sum_{H \ni e} g_i(H) = \sum_{e: f_i(e)>0} 1 \geqslant (1-\alpha)(m-t).$$

(Here the internal sum ranges over all graphs H in G_t that are isomorphic to K_k.) Since K_k has $\binom{k}{2}$ edges, this implies, in particular,

$$\tau_k^*(G_t) = v_k^*(G_t) \geqslant \frac{(1-\alpha)(m-t)}{\binom{k}{2}}.$$

By (3.3) we have

$$\tau_k^*(G) \geqslant \frac{(1-\alpha)(m-t)}{\binom{k}{2}} + \frac{t}{\lfloor k^2/4 \rfloor}. \tag{3.6}$$

The spanning subgraph P of G_t consisting of the edges having positive weight has $(1-\alpha)(m-t)$ edges. Since any graph can be made bipartite by removing less than half of its edges, we can delete from P a subset F of less than $(1-\alpha)(m-t)/2$ edges to make P bipartite.

We claim that the spanning subgraph Q of G_t obtained by removing F from G_t is K_k-free. Assume that Q has a K_k. The edges with positive weight form a bipartite subgraph on k vertices inside this K_k. The number of such edges is clearly at most $\lfloor k^2/4 \rfloor$. This implies that this K_k contains at least $\binom{k}{2} - \lfloor k^2/4 \rfloor$ edges with zero weight, contradicting the fact that G_t was the last graph in the sequence and has no copy of K_k with this number of zero weight edges. We have therefore proved

$$\tau_k(G_t) \leqslant \frac{(1-\alpha)(m-t)}{2}.$$

By (3.2) we have

$$\tau_k(G) \leqslant \frac{(1-\alpha)(m-t)}{2} + t. \tag{3.7}$$

By (3.6) and (3.7) we have

$$\tau_k^*(G) \geqslant \frac{2\tau_k(G) - 2t}{\binom{k}{2}} + \frac{t}{\lfloor k^2/4 \rfloor}. \tag{3.8}$$

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

So, (3.5) and (3.8) both supply lower bounds for $\tau_k^*(G)$ in terms of $\tau_k(G)$ and t. In particular, the maximum of both bounds can be used as a lower bound for $\tau_k^*(G)$. For $k \geqslant 4$ observe that (3.5) increases as t increases and (3.8) decreases as t increases. Hence, the maximum of both bounds is minimized when they are equal which, in turn, happens when $t = \tau_k(G)$. In this extremal point we have

$$\tau_k^*(G) \geqslant \frac{\tau_k(G)}{\lfloor k^2/4 \rfloor}.$$

Thus, $\tau_k(G) \leqslant \lfloor k^2/4 \rfloor \tau_k^*(G)$, proving the theorem. □

Theorem 3.2 immediately implies the following improvement of Corollary 3.1.

Corollary 3.3. $\tau_k(G) \leqslant \lfloor k^2/4 \rfloor \nu_k(G) + o(n^2)$.

It is well known that every graph with m edges can be made $(k-1)$-partite by removing from it less than $m/(k-1)$ edges. One just considers a random partition of the vertex set into $k-1$ parts and observes that the probability of an edge having both of its endpoints in the same part is less than $1/(k-1)$. In particular, $\tau_k(G) \leqslant m/(k-1) - o(m)$. As for the case of triangles, there are many different types of graphs that are hard to make K_k-free, that is, graphs for which $\tau_k(G) \geqslant m/(k-1) - o(m)$. We thus say that a graph G is $(1-\delta)$*-hard to make* K_k*-free* if $\tau_k(G) \geqslant (1-\delta)m/(k-1)$.

The following is an immediate consequence of Corollary 3.3.

Corollary 3.4. *Let G be a graph with m edges that is* $(1 - o_n(1))$*-hard to make* K_k*-free. Then,*

$$\nu_k(G) \geqslant \frac{m}{(k-1)\lfloor k^2/4 \rfloor} - o(n^2).$$

Observe that for, say, K_4 we get that dense graphs that are hard to make K_4-free have roughly $m/12$ edge-disjoint copies of K_4. As each K_4 has 6 edges, this implies that a fraction of roughly $1/2$ of the edges can be packed with edge-disjoint copies of K_4. More generally, for K_k, we get that a fraction of $2/k$ of the edges can be packed with edge-disjoint copies of K_k (if k is odd then this fraction is a bit larger). It is plausible that Conjecture 1.5 can be extended from triangles to larger cliques. That is, all but a negligible fraction of the edges can be packed with edge-disjoint copies of K_k.

4. Concluding remarks

The proof of Theorem 1.7 can be extended to other odd cycles. Denoting the edge-covering and packing numbers by $\tau_{C_k}(G)$ and $\nu_{C_k}(G)$ respectively, the analogous result states that for β-dense graphs that are $(1-\delta)$-hard to make C_k-free, we have

$$\nu_{C_k}(G) \geqslant (1 + c\beta)\frac{m}{2k-2},$$

where c is an absolute constant. Observe that for any graph G we have $\nu_{C_k}(G) \leqslant m/k$.

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

The proof is essentially the same, with the following minor differences. We use a straightforward extension of the result of Krivelevich for cycles of length k, which states that $\tau_{C_k}(G) \leqslant (k-1)\tau^*_{C_k}(G)$ (see also [4] for this observation), and the result of Haxell and Rödl applied to C_k stating that $|\nu_{C_k}(G) - \nu^*_{C_k}(G)| = o(n^2)$. As in the proof of Theorem 2.1, we split into three cases according to the relative sizes of F_0 and $F_{1/(k-2)}$, where the latter are all edges with weight at least $1/(k-2)$. Observe that this coincides with the definition of F_1 for the case of triangles. The only real difference is in Case 3. In Lemma 2.2 we can no longer claim that H is triangle-free. Rather, it is C_k-free. This means that any neighbourhood of a vertex is no longer forced to be an independent set, but rather it is forced not to contain a path of length $k-2$. But this, in turn, implies that each neighbourhood in H is sparse and has only a linear number of edges, which is negligible in the dense setting. Also, when using Lemma 2.2, by looking at the subgraph induced by $A \cup B$ in G, we can no longer claim that it contains only edges of $F_{1/(k-2)}$ with both endpoints in A or both endpoints in B. However, it certainly does not contain a path of length $k-2$ of edges not in $F_{1/(k-2)}$ with both endpoints in the same class. Thus, there are only a negligible (linear) number of edges not in $F_{1/(k-2)}$ that are inside A or inside B. Hence, the same argument as in Case 3 for triangles also holds here.

Theorem 3.2 supplies, in particular, an efficient approximation algorithm for the NP-hard problem of computing $\tau_k(G)$. Its approximation ratio is $\lfloor k^2/4 \rfloor$. It also bounds the integrality gap of this problem by $\lfloor k^2/4 \rfloor$.

Consider the problem of finding a maximal triangle-free subgraph. Its fractional relaxation is thus to assign weights in $[0,1]$ to the edges, so that for each triangle the sum of the weights is no larger than 2. The goal is to maximize the sum of the weights of such an assignment. Denoting the corresponding parameters by $\rho_3(G)$ and $\rho^*_3(G)$ we have, by definition, $\rho_3(G) = m - \tau_3(G)$ and $\rho^*_3(G) = m - \tau^*_3(G)$. The (asymptotic) integrality gap of this problem is known to be between 1.5 and 4/3. The lower bound comes from the complete graph. The integral solution is $n^2/4(1-o(1))$, while the fractional solution comes from assigning a weight of 2/3 to each edge, thereby obtaining a total weight of $n^2/3(1-o(1))$. The upper bound follows from Krivelevich's result $\tau_3(G) \leqslant 2\tau^*_3(G)$ after some easy arithmetic manipulations.

Our proof of Theorem 2.1 improves upon the upper bound for dense graphs. Suppose that G is a graph with $m = \beta n^2$ edges. Assume first that $\tau_3(G) \geqslant (1-\delta)m/2$. By Theorem 2.1, $\rho^*_3(G) \leqslant 3m/4 - m\beta/396$. On the other hand, for any graph we have $\rho_3(G) \geqslant m/2$. Thus, the integrality gap in this case is at most $3/2 - \beta/198$. Consider next the case $\tau_3(G) \leqslant (1-\delta)m/2$. Hence, $\rho_3(G) \geqslant (1+\delta)m/2$. By Krivelevich's result, we have

$$\rho_3(G) = m - \tau_3(G) \geqslant m - 2\tau^*_3(G) = 2\rho^*_3(G) - m.$$

This implies that the integrality gap is at most $1/2 + m/(2\rho_3(G))$. In our case this implies an integrality gap of $1/2 + 1/(1+\delta) = 3/2 - \delta/(1+\delta)$. Recall that Theorem 2.1 already holds for $\delta = \beta^2/99$.

Acknowledgement

I thank Guy Kortsarz and Michael Krivelevich for useful discussions and the referee for valuable comments. Special thanks are due to Wojciech Samotij, who pointed out that a small modification to the proof of Lemma 2.2 yields a linear dependence on β in the statement of Theorem 1.7 (the original proof had quadratic dependence).

https://doi.org/10.1017/S0963548312000235 Published online by Cambridge University Press

References

[1] Alon, N. (1996) Bipartite subgraphs. *Combinatorica* **16** 301–311.

[2] Haxell, P. (1999) Packing and covering triangles in graphs. *Discrete Math.* **195** 251–254.

[3] Haxell, P. E. and Rödl, V. (2001) Integer and fractional packings in dense graphs. *Combinatorica* **21** 13–38.

[4] Kortsarz, G., Langberg, M. and Nutov, Z. (2008) Approximating maximum subgraphs without short cycles. In *Proc. 11th International Workshop, APPROX 2008, and 12th International Workshop, RANDOM 2008 on Approximation, Randomization and Combinatorial Optimization*, Vol. 5171 of *Lecture Notes in Computer Science*, Springer, pp. 118–131.

[5] Krivelevich, M. (1995) On a conjecture of Tuza about packing and covering of triangles. *Discrete Math.* **142** 281–286.

[6] Tuza, Z. (1984) Conjecture. In *Finite and Infinite Sets: Eger, Hungary, 1981*, Vol. 37 of *Proc. Colloq. Math. Soc. Janos Bolyai*, North-Holland, p. 888.

[7] Yuster, R. (2005) Integer and fractional packing of families of graphs. *Random Struct. Alg.* **26** 110–118.